J. Dairy Sci. 101:1–7
https://doi.org/10.3168/jds.2018-14388

Short communication: Uncovering quantitative trait loci associated with resistance to *Mycobacterium avium* ssp. *paratuberculosis* infection in Holstein cattle using a high-density single nucleotide polymorphism panel

Sanjay Mallikarjunappa,*†‡ M. Sargolzaei,‡§ L. F. Brito,‡ K. G. Meade,† N. A. Karrow,‡ and S. D. Pant*[1]
*Graham Centre for Agricultural Innovation, NSW DPI and Charles Sturt University, Locked Bag 588, Wagga Wagga, NSW 2678, Australia
†Animal and Bioscience Research Department, Animal and Grassland Research and Innovation Centre, Teagasc, Grange, Co. Meath, Ireland, C15 PW93
‡Centre for Genetic Improvement of Livestock, Department of Animal Biosciences, University of Guelph, Guelph, N1G 2W1, Ontario, Canada
§The Semex Alliance, Guelph, Ontario, N1H 6J2, Canada

ABSTRACT

Mycobacterium avium ssp. *paratuberculosis* (MAP) is the etiological agent of Johne's disease in cattle. Johne's disease is a disease of significant economic, animal welfare, and public health concern around the globe. Therefore, understanding the genetic architecture of resistance to MAP infection has great relevance to advance genetic selection methods to breed more resistant animals. The objectives of this study were to perform a genome-wide association study of previously analyzed 50K genotypes now imputed to a high-density single nucleotide polymorphism panel (777K), aiming to validate previously reported associations and potentially identify additional single nucleotide polymorphisms associated with antibody response to MAP infection. A principal component regression-based genome-wide association study revealed 15 putative quantitative trait loci (QTL) associated with the MAP infection phenotype (serum or milk ELISA tests) on 9 different chromosomes (*Bos taurus* autosomes 5, 6, 7, 10, 14, 15, 16, 20, and 21). These results validated previous findings and identified new QTL on *Bos taurus* autosomes 15, 16, 20, and 21. The positional candidate genes *NLRP3*, *IFi47*, *TRIM41*, *TNFRSF18*, and *TNFRSF4* lying within these QTL were identified. Further functional validation of these genes is now warranted to investigate their roles in regulating the immune response and, consequently, cattle resistance to MAP infection.
Key words: Johne's disease, paratuberculosis, bovine high-density SNP panel, genome-wide association study

Short Communication

Johne's disease (**JD**) is a chronic intestinal inflammatory disease of ruminants caused by the gram-positive bacteria *Mycobacterium avium* ssp. *paratuberculosis* (**MAP**). With its worldwide distribution, JD causes major economic losses to the dairy industry due to decreased productivity in infected animals, premature culling, and management costs that are associated with controlling disease transmission (Ott et al., 1999; Coussens, 2001). Prevention of JD is greatly hindered by the lack of effective treatment options and high efficacy vaccines. However, resistance to MAP infection appears to be a heritable trait; thus, it may be possible to selectively breed animals for enhanced resistance to JD (Koets et al., 2000; Mortensen et al., 2004; Gonda et al., 2006; Küpper et al., 2012).

Pant et al. (2010) performed a genome-wide association study (**GWAS**), based on principal component regression analysis and Illumina BovineSNP50 (50K) genotypes, and identified QTL on BTA1, BTA5, BTA6, BTA7, BTA10, BTA11, and BTA14 that were associated with MAP infection status, as indicated by milk and serum antibody response to MAP infection (Pant et al., 2010). This study and other GWAS performed on different cattle populations suggest that resistance to JD is polygenic in nature (Gonda et al., 2007; Settles et al., 2009; van Hulzen et al., 2012; Alpay et al., 2014; Zare et al., 2014). However, there is a lack of congruence in the QTL identified, likely due to the different phenotypes that were used to define MAP infection status, the low heritable nature of these phenotypes, the complex nature of the disease progression, and the absence of QTL having a large effect on the disease phenotype (Kirkpatrick and Shook, 2011).

Previous GWAS have mostly relied on a 50K SNP panel to identify QTL associated with MAP infection

Received January 4, 2018.
Accepted March 29, 2018.
[1]Corresponding author: spant@csu.edu.au

status. Since infection status is influenced by a large number of genetic variants having small effects, it is certain that genotype data from such a low-density SNP panel will lack power to fine-map any causal variants. In recent years, with the improvement in genomic technologies and availability of high-density bovine 777K SNP panel (**HD**), it is now possible to perform GWAS aiming to identify QTL with increased accuracy (Erbe et al., 2012). Genotyping animals using HD panels is still expensive, but costs can be greatly reduced if genotyping is performed using a cheaper low-density (**LD**) SNP panel followed by genotype imputation to a HD SNP panel based on genotype information from a reference population of cattle.

Genotype imputation to HD SNP panel enables prediction of genotypes at loci that were not previously investigated using the LD platform. The imputed genotypes can then be tested for associations with a phenotype of interest by performing a GWAS, or be used to fine-map previously identified genomic regions harboring a QTL of interest (Marchini and Howie, 2010). Sargolzaei et al. (2014) previously demonstrated high imputation accuracy from a 50K SNP to HD SNP panel (777K), even with a small reference population of cattle.

The objectives of this study were to impute previously analyzed 50K data generated by Pant et al. (2010) to HD SNP panel (777K), and to perform a principal component regression-based GWAS on the imputed genotype data. The following analysis is based on the hypothesis that HD genotypes will enable identification of additional QTL associated with MAP infection status in cattle, especially in regions of low marker coverage on the previous 50K SNP platform.

Sample collection, classification of sampled animals as MAP positive (n = 90) and negative (n = 142) cohorts based on MAP infection status, and genotyping were previously described by Pant et al. (2010). Briefly, blood was collected from cows in 6 commercial herds in Southwestern and Eastern Ontario, Canada. Infection status of these animals was based on detection of MAP-specific antibodies in milk or blood plasma by ELISA test (IDEXX Laboratories, Westbrook, ME). Out of the 90 MAP-positive animals, 34 and 56 animals were considered as MAP-positive based on blood plasma and milk ELISA, respectively. Only the infection-free Holsteins that were older than 5.8 yr of age were chosen for the healthy (MAP negative) cohort (n = 142). Genomic DNA extracted from the buffy blood coat of a total of 232 animals were genotyped using the Illumina BovineSNP50 BeadChip (50K, Illumina Inc., San Diego, CA).

The 50K genotype data from 232 individuals was imputed to the BovineHD BeadChip SNP panel (777K, Illumina Inc.) using the FImpute software (Sargolzaei et al., 2014), based on the assumption that closely related individuals share longer haplotypes and distant relatives share shorter haplotypes. FImpute can perform imputation based on both family-based as well as population-based methods. As no pedigree information was available for our genotyped animals, imputation to the HD marker panel was performed using the population-based method.

Before performing GWAS, quality control (**QC**) measures were applied to imputed HD genotypes using the GENABEL package in R software (Aulchenko et al., 2007). Single-nucleotide polymorphisms with a minor allele frequency (**MAF**) lower than 10% were excluded. In total 522,197 SNP passed the QC filters and were considered for subsequent analyses.

A GWAS was performed on the imputed genotype data by 2-stage logistic regression approach as described in Pant et al. (2010). Briefly, single SNP logistic regression was performed in the first stage using the following model:

$$\text{Logit}\ (Y_i) = \mu + \beta\alpha + e_i,$$

where Y_i = phenotype of the animal (coded as 0 for MAP-negative and 1 for MAP-positive animals, respectively); μ = overall mean; β = regression coefficient; α = allele substitution effect coded as -1 for homozygote genotype mm, 0 for heterozygous genotype Mm or mM, and 1 for the other homozygous genotype MM; and e_i is the random error. In the second stage, all the significant SNP ($P < 0.05$) from the single SNP regression step were analyzed chromosome-wise by step-wise logistic regression as described in Pant et al. (2010) with the exception that the principal components (**PC**) were selected as covariates along with the SNP in question, which explained 80% of the total variance in the genomic relationship matrix (**G**) and all the duplicate genotype columns were removed before logistic regression. Backward selection based on Akaike's information criterion was used to drop PC terms that failed to improve the fit of the model. A complete description of the model as described by Pant et al. (2010) is as follows:

$$\text{Logit}\left(Y_i\right) = \mu + \beta_t\alpha_t + \sum_{j=1}^{k}\beta_j P_j + e_j,$$

where Y_i = binomial response phenotype of the ith animal; μ = overall mean; β_t = regression coefficient for the additive effect of the SNP_t; P_j = PC terms; β_j = multiple regression coefficients for the PC terms; and e_i = random error. The binomial response phenotype (dependent variable) was coded as in the preliminary

analysis. The coded coefficients for the additive (α_t) effect (independent variable) of SNP_t were as in the single SNP regression analysis. This procedure was repeated for each SNP on each chromosome, and *P*-values, odds ratios, and confidence intervals were obtained for all SNP. Multiple testing correction was applied using Šidák correction (Sidak, 1967) after the second stage of the analysis and was only based on the number of markers included in the second stage of our analysis.

Of the 522,197 SNP that passed QC filtering, 40,592 SNP were found to be significantly associated with antibody response to MAP at $P < 0.05$ in the first stage of single SNP logistic regression. The significant SNP were then analyzed chromosome-wise using PC multiple logistic regression model (second stage). The Sidak correction for the genome-wide multiple comparison threshold ($P < 1.99E\text{-}6$) was estimated based on 25,647 SNP (without duplicate genotypes). A total of 41 SNP across 9 chromosomes were found to be significant at $P < 1.99E\text{-}6$ (Table 1 and Figure 1). Chromosomal regions overlapping within 600 kb were assumed as single QTL, and based on this assumption, 41 SNP were grouped as 15 QTL found on BTA5, BTA6, BTA7, BTA10, BTA14, BTA15, BTA16, BTA20, and BTA21 (Table 1). A description of all the significant SNP, their accession numbers, chromosomal positions, odds ratio, their major and minor alleles, MAF, and *P*-value, along with gene symbols located within 1 Mb are shown in Table 1. In addition, the Manhattan plot for PC multiple regression analysis of HD imputed data is shown in Figure 1.

Other studies have also investigated the genetic architecture of MAP infection status in cattle. Significant genomic regions have been reported on various chromosomes such as BTA1 (Alpay et al., 2014), BTA2 (Alpay et al., 2014; Sallam et al., 2017), BTA3 (Settles et al., 2009; Zare et al., 2014; Sallam et al., 2017), BTA4 and BTA5 (van Hulzen et al., 2012), BTA6 (Alpay et al., 2014; Sallam et al., 2017), BTA7 (Alpay et al., 2014), BTA8 (Kiser et al., 2017; Sallam et al., 2017), BTA9 (Settles et al., 2009), BTA10, BTA12, and BTA14 (Kiser et al., 2017), BTA16 (Zare et al., 2014; Kiser et al., 2017), BTA17 (Alpay et al., 2014; Zare et al., 2014), BTA18 (van Hulzen et al., 2012), BTA20 (Gonda et al., 2007), BTA22 (Kiser et al., 2017), BTA23 (Zare et al., 2014), BTA25 and BTA27 (Sallam et al., 2017), BTA28 (van Hulzen et al., 2012), and BTA29 (Alpay et al., 2014; Sallam et al., 2017).

However, most of these studies were conducted using low- to moderate-density SNP panels that have lower genome coverage compared with the HD panel used in the current study (777K). Although genotyping animals with HD SNP panels is advantageous, it is also more expensive. Therefore, one alternative is to reanalyze LD genotypes by imputation to a HD SNP panel, which was the objective of the present study.

Principal component regression analysis of imputed HD genotypes revealed a total of 15 QTL on 9 different chromosomes to be associated with MAP infection in Holstein cattle. The list of gene symbols harboring the defined QTL regions is represented in Table 1. In the previous 50K genotype data analysis of the same resource population, Pant et al. (2010) reported QTL on BTA1, BTA5, BTA6, BTA7, BTA10, BTA11, and BTA14 to be significantly associated with MAP infection status. Associations of these QTL were validated in the current analysis on BTA5 and BTA10. Overlapping QTL on BTA7 and BTA14 were also observed when chromosome-wide threshold for multiple testing was used. However, it did not reach the significance threshold when the genome-wise correction was used. Although peaks were observed for QTL on BTA1 and BTA11, they did not cross the threshold of genome-wise and chromosome-wise multiple testing *P*-values. This could be due to the inclusion of PC that explained 80% of the total variance in the current analysis, as opposed to 90% as described in Pant et al. (2010). Another reason could be that different genome reference assemblies were used. The genome assembly used in this study was Bos_taurus_UMD_3.1.1, whereas BTAU 4.0 assembly was used by Pant et al. (2010) for annotation. As the bovine reference panels gets updated, there will be changes in the way genes are annotated, and Florea et al. (2011) reported major effects on gene content with the way genome sequences are assembled and updated. In addition to the overlapping QTL found in these studies, 7 additional QTL on BTA15, BTA16, BTA20, and BTA21 were identified in the present study. The identification of these additional QTL can be attributed to the higher number of SNP markers on 777K HD panel, which increase the power of GWAS.

Mapping of genes within the identified QTL regions revealed the presence of candidate genes that could be related to resistance to MAP infection in cattle. This includes genes such as *NLRP3*, *IFi47*, and *TRIM41* within the QTL defined on BTA7, and *TNFRSF18* and *TNFRSF4* on BTA16. *NLRP3* is a Nod-like intracellular innate immune receptor that recognizes pathogen-associated molecular patterns and triggers induction of pro-inflammatory cytokines IL-1β and IL-18 through caspase-1 activation within the inflammasome complex (Zaki et al., 2011). Interestingly, SNP on *NLRP3* region contribute to Crohn's disease susceptibility in humans (Villani et al., 2009), and MAP has been linked with this disease (Scanu et al., 2007). The second candidate gene *IFi47*, interferon gamma inducible protein 47 or *IRG47*,

Table 1. A description of the SNP significantly associated with *Mycobacterium avium* ssp. *paratuberculosis* infection status[1]

QTL	SNP rsID	BTA	Position (bp)	MAF	Odds ratio	*P*-value	Gene symbols (within 1 Mbp)	Overlap
1	rs135242518	5	81,530,672	0.27	0.11786	3.87E-07	*CCDC91*, *PTHLH*, *FAR2*, *MANSC*, *REP15*, *PPFIBP1*	Yes
	rs109339854	5	82,482,567	0.48	0.08545	2.65E-08	*ERGIC2*, *KLHDC5*, *MANSC4*, *MRPS35*	
	rs134245823	5	82,494,874	0.48	11.7032	2.65E-08	*ARNTL2*, *STK38L*, *MED21*, *TM7SF3*, *FGFR1OP2*, *ASUN*, *ITPR2*	
2	rs134810456	6	19,064,492	0.3	24.3705	8.52E-10	*PAPSS1*, *DKK2*, *SGMS2*, *CYP2U1*, *HADH*, *LEF1*	No
3	rs132756399	7	41,974,204	0.43	0.04368	1.44E-06	*NLRP3*, *IFI47*, *OR2B11*, *TRIM52*	No
	rs110601314	7	41,980,231	0.43	22.4152	1.35E-06	*GNB2L1*, *TRIM41*, *OR2C3*, *TRIM7*, *OR2G2*, *OR2V2*	
	rs110450821	7	41,983,705	0.43	0.04461	1.35E-06	*BTNL9*, *COX7B*, *OR2G3*	
4	rs42556867	10	50,670,173	0.41	5.73785	2.45E-07	*GTF2A2*, *BNIP2*, *GCNT3*, *FAM81A*, *MYO1E*	Yes
	rs134355633	10	50,679,466	0.41	0.17428	2.45E-07	*FOXB1*, *CCNB2*, *RNF111*, *ANXA2*, *FAM63B*	
	rs42556851	10	50,705,459	0.41	0.15636	7.72E-08	*NARG2*, *RORA*, *ADAM10*, *AQP9*, *ALDH1A2*	
	rs133988414	10	50,921,608	0.4	6.0082	1.24E-07	*POLR2M*	
	rs132817923	10	50,971,153	0.4	0.16644	1.24E-07		
	rs43629042	10	52,023,061	0.35	0.16201	2.90E-08		
5	rs132870210	10	59,371,041	0.44	0.17253	2.43E-07	*CYP19A1*, *AP4E1*, *GLDN*, *DMXL2*, *SPPL2A*, *TRPM7*, *SCG3*, *USP50*, *USP8*, *LYSMD2*, *TMOD2*, *GABPB1*, *HDC*, *TMOD3*, *LEO1*, *SLC27A2*, *MAPK6*, *TBPL2*	Yes
6	rs110242629	10	68,119,052	0.41	9.19329	6.18E-09	*TBPL2*, *ATG14*, *FBXO34*, *KTN1*, *DLGAP5*, *LGALS3*, *MAPK1IP1L*, *SOCS4*, *WDHD1*, *GCH1*, *SAMD4A*, *PELI2*, *MIR2292*, *CGRRF1*, *GMFB*, *CNIH*, *CDKN3*	Yes
7	rs110813155	14	48,158,499	0.47	0.12589	9.09E-09	*SAMD12*, *MIR2489*, *EXT1*, *TNFRSF11B*, *MED30*, *COLEC10*, *MAL2*	No
	rs110774927	14	48,161,079	0.47	7.94325	9.09E-09		
8	rs41742892	14	49,965,170	0.28	0.16089	5.50E-08	*EIF3H*, *UTP23*, *RAD21*, *AARD*, *TRPS1*	No
9	rs135496866	15	24,778,751	0.23	0.08457	1.71E-07	*HTR3B*, *USP28*, *HTR3A*, *CLDN25*	No
	rs41748278	15	24,783,744	0.18	13.4531	1.13E-07	*ZW10*, *ZBTB16*, *TMPRSS5*, *DRD2*	
	rs42767091	15	24,792,798	0.23	11.8242	1.71E-07	*RBM7*, *REXO2*, *TTC12*, *NCAM1*, *NXPE4*	
10	rs134945484	16	49,006,258	0.32	5.05355	3.63E-07	*TNNT2*, *LAD1*, *TNNI1*, *CSRP1*, *IPO9*	No
	rs132976617	16	49,009,928	0.32	0.19788	3.63E-07	*TIMM17A*, *SHISA4*, *MIR2320*,	
	rs137767503	16	49,886,643	0.38	0.1844	1.53E-07	*TPRG1L*, *MIR551A*, *ARHGEF16*, *PRDM16*	
	rs133093887	16	49,896,349	0.38	1.5E-07	1.53E-07		
11	rs136182707	16	53,247,138	0.41	0.1155	3.95E-08	*TNFRSF18*, *TNFRSF4*, *DNAJC16*, *CASP9*, *CELA2A*, *AGMAT*, *CTRC*	No
	rs133196443	16	53,247,913	0.41	8.65791	3.95E-08	*EFHD2*, *PLEKHM2*, *SLC25A34*, *TMEM82*, *FBLIM1*, *SPEN*, *TMEM51*	
	rs110621545	16	53,264,276	0.44	0.14361	2.90E-08	*ZBTB17*, *KAZN*, *SAMD11*, *NOC2L*, *KLHL17*, *C16H1orf170*, *HES4*	
	rs209889900	16	53,265,578	0.46	7.2721	2.21E-08	*ISG15*, *AGRN*, *RNF223*, *C16H1orf159*, *ACAP3*, *SCNN1D*, *PUSL1*	
	rs41811150	16	53,277,801	0.33	0.13802	1.02E-08	*CPSF3L*, *GLTPD1*, *MIR200B*, *C16H1orf233*, *MIR200A*, *MIR429*	
	rs41810207	16	53,288,654	0.33	7.24557	1.02E-08	*AURKAIP1*, *CCNL2*, *MRPL20*, *ANKRD65*, *SDF4*, *B3GALT6*, *FAM132A*, *UBE2J2*, *DVL1*, *MXRA8TMEM88B*, *VWA1*, *ATAD3A*	
12	rs41810662	16	55,677,310	0.34	0.16549	1.29E-07	*PRAMEF12*, *CYB5R1*, *ADIPOR1*, *KLHL12*, *RABIF*, *ZBTB37*	No
	rs43040872	16	55,721,889	0.34	0.15434	4.27E-08	*PDPN*, *KDM5B*, *C16H1orf158*, *PRDM2*, *PRDX6*, *RABGAP1L*	
	rs134792861	16	55,722,841	0.33	0.1641	8.03E-08	*SLC9C2*, *ANKRD45*, *KLHL20*, *CENPL*, *DARS2*, *SERPINC1*	
	rs109119538	16	55,839,833	0.42	0.13217	5.82E-08		
	rs135066039	16	56,108,258	0.35	5.63227	2.76E-07		
	rs42023278	16	56,114,387	0.34	0.13958	2.96E-08		
	rs43739343	16	56,120,577	0.35	0.17044	1.91E-07		
13	rs110473149	20	29,320,129	0.47	4.62079	3.71E-07	*HCN1*, *EMB*, *MRPS30*, *PARP8*	No
14	rs136813055	20	24,527,108	0.31	0.2112	9.90E-07	*SNX18*, *HSPB3*, *ESM1*, *GZMK*, *ARL15*, *GZMA*, *CDC20B*, *GPX8*, *MIR449A*, *MCIDAS*, *CCNO*, *DHX29*, *PPAP2A*, *SLC38A9*, *NDUFS4*	No
15	rs133273791	21	71,338,809	0.16	12.6771	6.17E-07	*PACS2*, *CRIP1*, *C21H14orf80*, *BRF1*, *BTBD6*, *NUDT14*, *JAG2*, *GPR132*, *CDCA4*, *C21H14orf79*, *AHNAK2*, *PLD4*, *KIAA0284*, *ZBTB42*, *SIVA1*, *ADSSL1*, *INF2*, *TMEM179*, *NRAC*	No

[1]MAF = minor allele frequency; overlap = gene overlap with previous 50K analysis.

codes for a GTPase that regulates IFN-γ-dependent macrophage resistance against intracellular bacteria and protozoa (Taylor et al., 2004). The third candidate gene (*TRIM41*) codes for a member of the tripartite motif family that has recently been characterized as a *NOD2*-interacting protein that may be involved in the *NOD2* cell signaling pathway (Thiébaut et al., 2016). The *NOD2* (*CARD 15*) is well known to have a protective function. The *NOD2* is a pathogen recognition receptor that recognizes the mycobacterial pathogen-associated molecular pattern muramyl dipeptide, which leads to activation of transcription factor NF-κB and pro-inflammatory cytokine expression (Girardin et al., 2003; Abbott et al., 2004), and SNP in *NOD2* have previously been associated with susceptibility to MAP infection in cattle (Ruiz-Larrañaga et al., 2010). The fourth and fifth candidate genes (*TNFRSF18* and *TNFRSF4*) belong to the tumor necrosis factor receptor super family and are known to stimulate a T-helper 1 cell-mediated immune response, which is important for controlling MAP infection, and to regulate the inflammatory response (Hehlgans and Pfeffer, 2005).

A final interesting gene harboring a QTL on BTA15 is also worth mentioning because of its immunological importance; the *IL-18* gene is located at 1.9 Mb upstream from the SNP rs135496866. The protective pro-inflammatory role of *IL-18* against mycobacterial infections is well documented. For instance, Alfonseca-Silva et al. (2016) reported increased expression of *IL-18* from *Mycobacterium bovis*-infected macrophages that has been isolated from naturally resistant as opposed to susceptible donors. Sugawara et al. (1999) also demonstrated the protective role of *IL-18* through IFN-γ induction during the course of mycobacterial infection.

In comparison with this study, 2 recently published GWAS identified different positional candidate genes and QTL associated with JD. By conducting a combined across-breed GWAS, Sallam et al. (2017) identified QTL on BTA2, BTA3, BTA6, BTA8, BTA25, BTA27, and BTA29 and also reported 2 positional candidate genes, *BTN1A1* and *TDP2*, associated with JD (cases were defined as cows that were positive by either fecal culture or serum ELISA tests). In an another GWAS using BovineSNP50 data, where JD cases were defined as MAP tissue quantitative PCR and culture positive, Kiser et al. (2017) identified 7 QTL located on BTA22 in Jerseys and 6 QTL located on BTA8, BTA10, BTA12, BTA14, and BTA16 in the combined Pacific Northwest and Northeast Holstein population by meta-analysis. The positional candidate genes identified in this study included *BCAR3*, *FLVCR2*, *RASA3*, *MGC134473*, *MARK1*, *C16H1orf115*, *MARC2*, *C10H14ORF1*, and *CDC42BPA*. All these genes play an important role in immunological processes, with some potentially affecting MAP entry into host immune cells. The dissimilarity in the findings of these 2 studies compared with the current study can be attributed to different phenotypes used to define JD positive animals, different statistical analysis employed for the GWAS, different sample size, and different cattle populations. It is of great importance to perform GWAS across independent populations and breeds to validate previous findings and better define the genetic architecture of JD, as well as identify important genomic regions affecting resistance to JD across breeds and populations. The identification of dif-

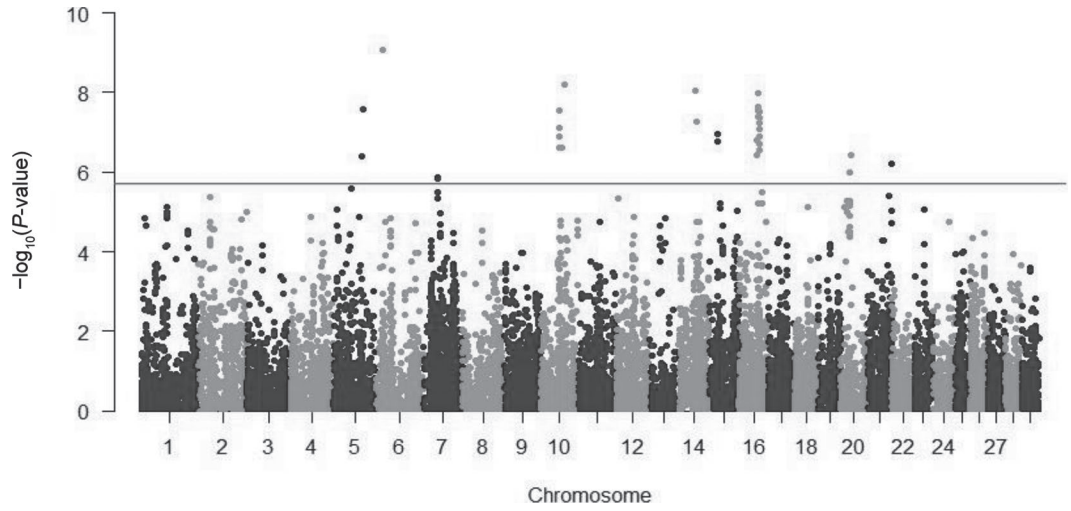

Figure 1. Manhattan plot for principal component regression analysis of imputed genotype data. A total of 41 SNP across 9 chromosomes (BTA5, BTA6, BTA7, BTA10, BTA14, BTA15, BTA16, BTA20, and BTA21) were found to be significant at $P < 1.99E\text{-}6$ (Sidak correction). Color version available online.

ferent QTL and candidate genes across chromosomes and populations is a clear reflection of the polygenic nature of JD.

Limitations associated with this study include the small size of the resource population (n = 232) and the lack of pedigree records. However, the 2-stage logistic regression method used for this GWAS allowed for joint analyses of individual SNP, while also using PC from the SNP-covariance matrix as regressors to account for linkage disequilibrium that exists between markers present on the same chromosome. Simultaneous analysis of multiple markers accounted for the population substructure, which was essential because the pedigree details of the case-control animals were unknown. Principal components were also only computed for markers that were found to be significant after the first stage of the analysis, which avoided infinite likelihood and inaccurate estimates of the phenotype. The multiple logistic PC regression analysis and its advantages over traditional single-SNP regression analysis are further explained in detail by Pant et al. (2010).

The genotype imputation used in the present study provided an inexpensive method for predicting genotypes and identifying QTL associated with MAP infection status. Imputation accuracy using FImpute software, which is mainly influenced by the density of SNP panel and the number of individuals making up the reference population (n = 2,998), was expected to be greater than 98% in the present study (Sargolzaei et al., 2014). These authors have also shown that imputation accuracy for rare alleles having a MAF <0.05 is higher using FImpute in comparison to other imputation software such as Beagle and Impute2.

In conclusion, a GWAS was performed using HD-imputed genotype data to identify QTL associated with MAP infection status. Several new QTL harboring potential candidate genes were identified that could be involved in the immune response to MAP infection. These associations should be validated in an independent population, and the functional contribution of these genes in the host response to MAP warrants further investigation.

ACKNOWLEDGMENTS

We thank the following funding organizations for supporting this work: The Semex Alliance, Guelph, ON, Canada; Natural Sciences and Engineering Research Council of Canada (NSERC); the Graham Centre for Agricultural Innovation, Wagga Wagga, NSW, Australia; Darcy John O'Sullivan Bequest, School of Animal and Veterinary Sciences, Charles Sturt University; and the Teagasc Walsh Fellowship Award to SM.

REFERENCES

Abbott, D. W., A. Wilkins, J. M. Asara, and L. C. Cantley. 2004. The Crohn's disease protein, NOD2, requires RIP2 in order to induce ubiquitinylation of a novel site on NEMO. Curr. Biol. 14:2217–2227. https://doi.org/10.1016/j.cub.2004.12.032.

Alfonseca-Silva, E., R. Hernández-Pando, and J. A. Gutiérrez-Pabello. 2016. *Mycobacterium bovis*-infected macrophages from resistant and susceptible cattle exhibited a differential pro-inflammatory gene expression profile depending on strain virulence. Vet. Immunol. Immunopathol. 176:34–43. https://doi.org/10.1016/j.vetimm.2016.02.015.

Alpay, F., Y. Zare, M. H. Kamalludin, X. Huang, X. Shi, G. E. Shook, M. T. Collins, and B. W. Kirkpatrick. 2014. Genome-wide association study of susceptibility to infection by *Mycobacterium avium* subspecies *paratuberculosis* in Holstein cattle. PLoS One 9:e111704. https://doi.org/10.1371/journal.pone.0111704.

Aulchenko, Y. S., S. Ripke, A. Isaacs, and C. M. van Duijn. 2007. GenABEL: An R package for genome-wide association analysis. Bioinformatics 23:1294–1296.

Coussens, P. M. 2001. *Mycobacterium paratuberculosis* and the bovine immune system. Anim. Health Res. Rev. 2:141–161.

Erbe, M., B. J. Hayes, L. K. Matukumalli, S. Goswami, P. J. Bowman, C. M. Reich, B. A. Mason, and M. E. Goddard. 2012. Improving accuracy of genomic predictions within and between dairy cattle breeds with imputed high-density single nucleotide polymorphism panels. J. Dairy Sci. 95:4114–4129. https://doi.org/10.3168/jds.2011-5019.

Florea, L., A. Souvorov, T. S. Kalbfleisch, and S. L. Salzberg. 2011. Genome assembly has a major impact on gene content: A comparison of annotation in two *Bos taurus* assemblies. PLoS One 6. https://doi.org/10.1371/journal.pone.0021400.

Girardin, S. E., J. P. Hugot, and P. J. Sansonetti. 2003. Lessons from Nod2 studies: Towards a link between Crohn's disease and bacterial sensing. Trends Immunol. 24:652–658. https://doi.org/10.1016/j.it.2003.10.007.

Gonda, M. G., Y. M. Chang, G. E. Shook, M. T. Collins, and B. W. Kirkpatrick. 2006. Genetic variation of *Mycobacterium avium* ssp. *paratuberculosis* infection in US Holsteins. J. Dairy Sci. 89:1804–1812. https://doi.org/10.3168/jds.S0022-0302(06)72249-4.

Gonda, M. G., B. W. Kirkpatrick, G. E. Shook, and M. T. Collins. 2007. Identification of a QTL on BTA20 affecting susceptibility to *Mycobacterium avium* ssp. *paratuberculosis* infection in US Holsteins. Anim. Genet. 38:389–396. https://doi.org/10.1111/j.1365-2052.2007.01627.x.

Hehlgans, T., and K. Pfeffer. 2005. The intriguing biology of the tumour necrosis factor/tumour necrosis factor receptor superfamily: players, rules and the games. Immunology 115:1–20. https://doi.org/10.1111/j.1365-2567.2005.02143.x.

Kirkpatrick, B. W., and G. E. Shook. 2011. Genetic susceptibility to paratuberculosis. Vet. Clin. North Am. Food Anim. Pract. 27:559–571. https://doi.org/10.1016/j.cvfa.2011.07.003.

Kiser, J. N., S. N. White, K. A. Johnson, J. L. Hoff, J. F. Taylor, and H. L. Neibergs. 2017. Identification of loci associated with susceptibility to *Mycobacterium avium* subspecies *paratuberculosis* (MAP) tissue infection in cattle. J. Anim. Sci. 95:1080–1091. https://doi.org/10.2527/jas.2016.1152.

Koets, A. P., G. Adugna, L. L. Janss, H. J. van Weering, C. H. Kalis, G. H. Wentink, V. P. Rutten, and Y. H. Schukken. 2000. Genetic variation of susceptibility to *Mycobacterium avium* subsp. *paratuberculosis* infection in dairy cattle. J. Dairy Sci. 83:2702–2708. https://doi.org/10.3168/jds.S0022-0302(00)75164-2.

Küpper, J., H. Brandt, K. Donat, and G. Erhardt. 2012. Heritability estimates for *Mycobacterium avium* subspecies *paratuberculosis* status of German Holstein cows tested by fecal culture. J. Dairy Sci. 95:2734–2739. https://doi.org/10.3168/jds.2011-4994.

Marchini, J., and B. Howie. 2010. Genotype imputation for genome-wide association studies. Nat. Rev. Genet. 11:499–511. https://doi.org/10.1038/nrg2796.

Mortensen, H., S. S. Nielsen, and P. Berg. 2004. Genetic variation and heritability of the antibody response to *Mycobacterium avium*

subspecies *paratuberculosis* in Danish Holstein cows. J. Dairy Sci. 87:2108–2113. https://doi.org/10.3168/jds.S0022-0302(04)70029-6.

Ott, S. L., S. J. Wells, and B. A. Wagner. 1999. Herd-level economic losses associated with Johne's disease on US dairy operations. Prev. Vet. Med. 40:179–192. https://doi.org/10.1016/S0167-5877(99)00037-9.

Pant, S. D., F. S. Schenkel, C. P. Verschoor, Q. You, D. F. Kelton, S. S. Moore, and N. A. Karrow. 2010. A principal component regression based genome wide analysis approach reveals the presence of a novel QTL on BTA7 for MAP resistance in Holstein cattle. Genomics 95:176–182. https://doi.org/10.1016/j.ygeno.2010.01.001.

Ruiz-Larrañaga, O., J. M. Garrido, M. Iriondo, C. Manzano, E. Molina, A. P. Koets, V. P. M. G. Rutten, R. A. Juste, and A. Estonba. 2010. Genetic association between bovine NOD2 polymorphisms and infection by *Mycobacterium avium* ssp. *paratuberculosis* in Holstein-Friesian cattle. Anim. Genet. 41:652–655. https://doi.org/10.1111/j.1365-2052.2010.02055.x.

Sallam, A. M., Y. Zare, F. Alpay, G. E. Shook, M. T. Collins, S. Alsheikh, M. Sharaby, and B. W. Kirkpatrick. 2017. An across-breed genome wide association analysis of susceptibility to paratuberculosis in dairy cattle. J. Dairy Res. 84:61–67. https://doi.org/10.1017/S0022029916000807.

Sargolzaei, M., J. P. Chesnais, and F. S. Schenkel. 2014. A new approach for efficient genotype imputation using information from relatives. BMC Genomics 15:478. https://doi.org/10.1186/1471-2164-15-478.

Scanu, A. M., T. J. Bull, S. Cannas, J. D. Sanderson, L. A. Sechi, G. Dettori, S. Zanetti, and J. Hermon-Taylor. 2007. *Mycobacterium avium* subspecies *paratuberculosis* infection in cases of irritable bowel syndrome and comparison with Crohn's disease and Johne's disease: Common neural and immune pathogenicities. J. Clin. Microbiol. 45:3883–3890. https://doi.org/10.1128/JCM.01371-07.

Settles, M., R. Zanella, S. D. McKay, R. D. Schnabel, J. F. Taylor, R. Whitlock, Y. Schukken, J. S. Van Kessel, J. M. Smith, and H. Neibergs. 2009. A whole-genome association analysis identifies loci associated with *Mycobacterium avium* ssp. *paratuberculosis* infection status in US Holstein cattle. Anim. Genet. 40:655–662. https://doi.org/10.1111/j.1365-2052.2009.01896.x.

Sidak, M. 1967. Rectangular confidence regions for the means of multivariate normal distributions. J. Am. Stat. Assoc. 62:626–633.

Sugawara, I., H. Yamada, H. Kaneko, S. Mizuno, K. Takeda, and S. Akira. 1999. Role of interleukin-18 (*IL-18*) in mycobacterial infection in *IL-18*-gene-disrupted mice. Infect. Immun. 67:2585–2589.

Sweeney, R. W. 1996. Transmission of paratuberculosis. The veterinary clinics of North America. Food Anim. Pract. 12:305–312. https://doi.org/10.1016/S0749-0720(15)30408-4.

Taylor, G. A., C. G. Feng, and A. Sher. 2004. p47 GTPases: Regulators of immunity to intracellular pathogens. Nat. Rev. Immunol. 4:100–109. https://doi.org/10.1038/nri1270.

Thiébaut, R., S. Esmiol, P. Lecine, B. Mahfouz, A. Hermant, C. Nicoletti, S. Parnis, J. Perroy, J. P. Borg, L. Pascoe, J. P. Hugot, and V. Ollendorff. 2016. Characterization and genetic analyses of new genes coding for NOD2 interacting proteins. PLoS One 11. https://doi.org/10.1371/journal.pone.0165420.

van Hulzen, K. J. E., A. P. Koets, M. Nielen, J. Hoeboer, J. A. M. van Arendonk, and H. C. M. Heuven. 2012. Genetic variation for infection status as determined by a specific antibody response against *Mycobacterium avium* subspecies *paratuberculosis* in milk of Dutch dairy goats. J. Dairy Sci. 95:6145–6151. https://doi.org/10.3168/jds.2012-5616.

Villani, A.-C., M. Lemire, G. Fortin, E. Louis, M. S. Silverberg, C. Collette, N. Baba, C. Libioulle, J. Belaiche, A. Bitton, D. Gaudet, A. Cohen, D. Langelier, P. R. Fortin, J. E. Wither, M. Sarfati, P. Rutgeerts, J. D. Rioux, S. Vermeire, T. J. Hudson, and D. Franchimont. 2009. Common variants in the *NLRP3* region contribute to Crohn's disease susceptibility. Nat. Genet. 41:71–76. https://doi.org/10.1038/ng.285.

Zaki, M. H., M. Lamkanfi, and T. D. Kanneganti. 2011. The Nlrp3 inflammasome: Contributions to intestinal homeostasis. Trends Immunol. 32:171–179. https://doi.org/10.1016/j.it.2011.02.002.

Zare, Y., G. E. Shook, M. T. Collins, and B. W. Kirkpatrick. 2014. Genome-wide association analysis and genomic prediction of *Mycobacterium avium* subspecies *paratuberculosis* infection in US Jersey cattle. PLoS One 9:e88380. https://doi.org/10.1371/journal.pone.0088380.